



# TERRA POPULUS

A Global Population/Environment Data Network



# The explosion of Scientific Data

*Because of the massive decline in the cost of data collection, storage, and analysis, the quantity of scientific data being collected is growing at an extraordinary pace*

- New opportunities for analysis
- New methods are being applied
- Marked acceleration in the pace of discovery



# The Big Challenges

*The quantity of scientific data is exploding, but we lack basic infrastructure to maintain them or capitalize on opportunities for analysis and discovery*

- Most scientific data is at risk of loss
- Most scientific data is inaccessible
- Metadata are usually incomplete and inadequate
- Little interoperability across datasets or data types
- Data are trapped in disciplinary silos



# TerraPop Goals

Provide an organizational and technical framework to preserve, integrate, and disseminate global-scale spatiotemporal data describing population and the environment.

- Census microdata
- Government land-use statistics
- Land cover data from satellite imagery
- Historical climate records (temperature, precipitation, cloud cover)



Age  
Sex  
Relationship  
Race  
Birthplace  
Mother's birthplace  
Occupation

H9100002400000000088001001000220100
P910000020101032120010010010011504
P910000010201036220010010010011999
P910201000301011220060010010011999
P910201000301009120060010010011999
P910201000301007120060010010011999
P910201000301006120060010010011999
P910201000301004220060010010011999
P910201000301003220060010010011999
P910201000301002220060010010011999
H9100002400000000088001001000110100
P910000020101030110010290510511310
P910000010201021210010290290171999
P910201000301001110060010290291999
H9100002400000000088001001000220100
P910000020101045120010010010011100
P910000010201025220010010010011820
P910201000301007220060010010011999
H9100002400000000088001001000220100
P910000020101049120010010010011100
P910000010201049220010010010011820
P910201000301019220060010010011820
P910201000301015220060010010012820

## Microdata Structure

Geographic and housing characteristics

Household record (shaded) followed by a person record for each member of the household

For each type of record, columns correspond to specific variables



# The Power of Microdata

- **Customized measures:** Variables based on combined characteristics of family and household members, capitalizing on the hierarchical structure of the data
- **Multivariate analysis:** Analyze many individual, household, and community characteristics simultaneously
- **Interoperability:** Harmonize data across time and space

For each person, detailed information about geographic location, economic activities, educational attainment, literacy, fertility history, child mortality, migration, place of former residence, marital status, consensual unions, family composition, disabilities, water supply, sewage, building materials (floor, roof, etc.), and many other characteristics.





[Home](#) | [Select Data](#) | [FAQ](#) | [Contact](#) | [Login](#)

## PROJECT

About IPUMS-I  
How to Cite IPUMS-I  
User Registration and Login

## DATA

Browse and Select Data  
Download Your Data Extract  
GIS and Other Data Files

## SAMPLES

Sample Descriptions  
Variance Estimation  
Source Documents

## RESOURCES

International Partners  
World Data Inventory  
Microdata Handbook  
Bibliography

# Integrated Public Use Microdata Series, International

census microdata for social and economic research

IPUMS-International is a project dedicated to collecting and distributing census data from around the world. Its goals are to:

- Collect and preserve data and documentation
- Harmonize data
- Disseminate the data absolutely free!

62 countries - 185 censuses - 397 million person records

## IPUMSI News

June 2011 data release  
2010 award winners  
Improved web interface  
IPUMS Havana workshop  
June 2010 data release  
Mortality and fertility data  
NIH extends IPUMS-I  
... All news items

## MPC Data Projects

[IPUMS-USA and others](#)



# NAPP

## North Atlantic Population Project

[Home](#) [Select Data](#) [FAQ](#) [Contact](#) [Login](#) [Data Cart](#)

### PROJECT

[About NAPP](#)  
[Data Releases](#)  
[Revision History](#)  
[User Registration & Login](#)

### DATA

[Browse and Select Data](#)  
[Download Data Extract](#)  
[Linked Samples](#)  
[Online Data Analysis](#)

### DOCUMENTATION

[Samples](#)  
[Census Questionnaires](#)  
[Other Documentation](#)

### RESOURCES

[NAPP Participants](#)  
[Citation and Use](#)  
[Bibliography](#)

## North Atlantic Population Project

Census microdata from Canada, Great Britain, Germany, Iceland, Norway, Sweden, and the United States from 1801 to 1910. The project's goals are to:

- Harmonize data, including many complete count datasets
- Link individuals between census years for longitudinal analysis
- Disseminate the data absolutely free!



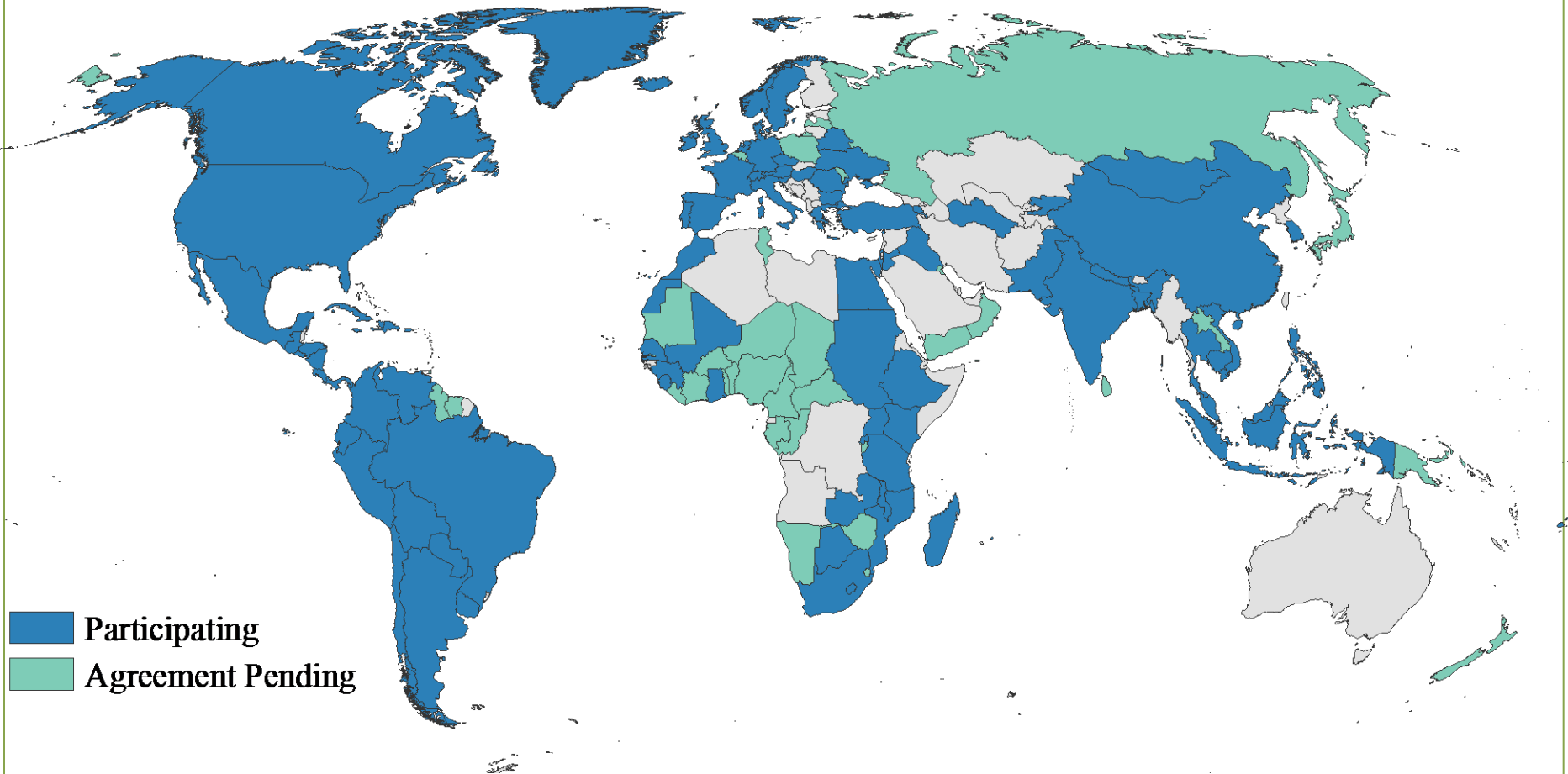
### NAPP News

**NEW!** Full-count data for Iceland 1801 and 1901  
**NEW!** Full-count data for Norway 1801  
**NEW!** Samples for Canada 1852 and 1891  
[Linked data for Norway and the U.S.](#)  
[Mecklenburg-Schwerin 1819 sample](#)

### Other MPC Projects

[IPUMS-International](#)  
[IPUMS-USA](#)  
[IPUMS-CPS](#)  
[IHIS](#)  
[NHGIS](#)

# IPUMS/NAPP Participating Countries



# IPUMS/NAPP Data Releases

2011:

- 600 Million persons
- 300 censuses and surveys
- 65 countries

2016:

- 1.2 billion persons
- 800 censuses and surveys
- 110 countries



# TerraPop Partners



INSTITUTE ON THE  
ENVIRONMENT

UNIVERSITY OF MINNESOTA  
Driven to Discover<sup>SM</sup>



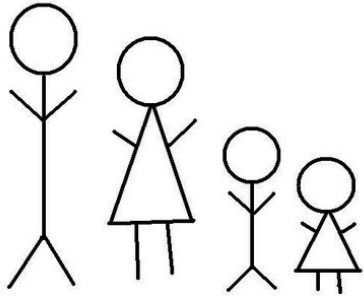
ICPSR | INTER-UNIVERSITY  
CONSORTIUM FOR  
POLITICAL AND  
SOCIAL RESEARCH



UNIVERSITY OF MINNESOTA  
LIBRARIES

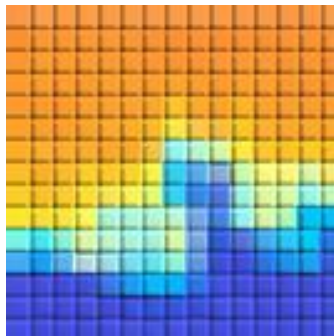
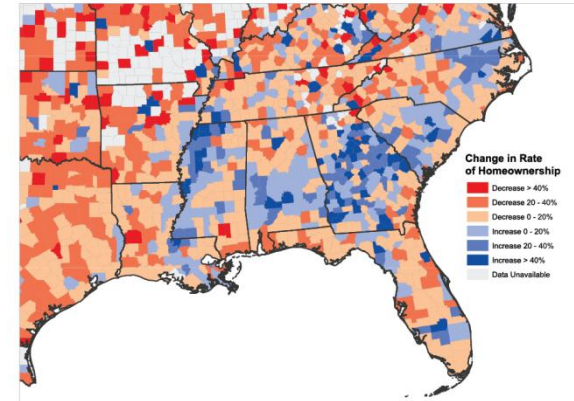


# Three data types



**Microdata:**  
Characteristics of individuals  
and households

**Small-area data:**  
Characteristics of places defined  
by administrative boundaries



**Gridded data:**  
Values arranged in rows  
and columns



# Three output formats

1. Census microdata with attached characteristics describing land use, land cover, and climate for local areas
2. Aggregate data for administrative districts with tabulated population data and environmental characteristics
3. Gridded data with characteristics of population and environment

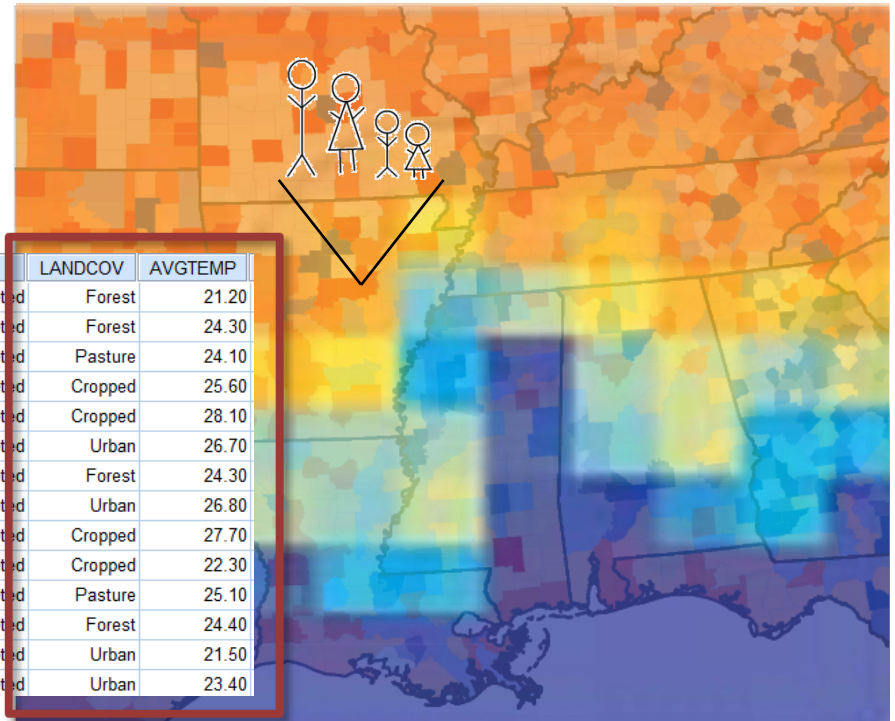


# Data Fusion – Microdata Output

Census microdata with attached characteristics describing land use, land cover, and climate for local areas

Individuals and households  
with their environmental  
and social context

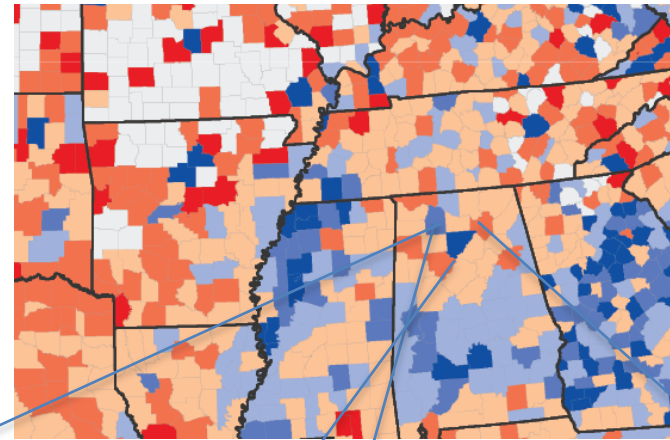
YEAR	AGE	SEX	MGRATE5	EDATTAN	LANDCOV	AVGTEMP
1991	10	Male	Same major, same minor administrative unit	Less than primary completed	Forest	21.20
1991	27	Female	Same major, different minor administrative unit	Secondary completed	Forest	24.30
1991	54	Female	Same major, same minor administrative unit	Primary completed	Pasture	24.10
1991	37	Male	Same major, same minor administrative unit	University completed	Cropped	25.60
1991	37	Female	Same major, same minor administrative unit	University completed	Cropped	28.10
1991	42	Female	Different major administrative unit	Less than primary completed	Urban	26.70
1991	20	Female	Different major administrative unit	Less than primary completed	Forest	24.30
1991	39	Male	Same major, same minor administrative unit	University completed	Urban	26.80
1991	77	Female	Same major, same minor administrative unit	Less than primary completed	Cropped	27.70
1991	11	Female	Same major, same minor administrative unit	Less than primary completed	Cropped	22.30
1991	31	Female	Same major, same minor administrative unit	University completed	Pasture	25.10
1991	23	Male	Same major, same minor administrative unit	Primary completed	Forest	24.40
1991	24	Female	Same major, same minor administrative unit	University completed	Urban	21.50
1991	40	Female	Same major, same minor administrative unit	University completed	Urban	23.40



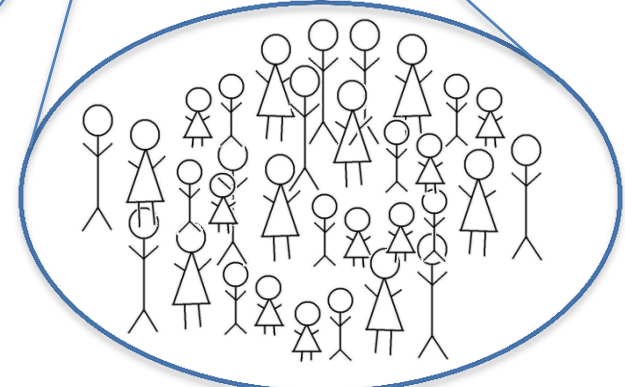
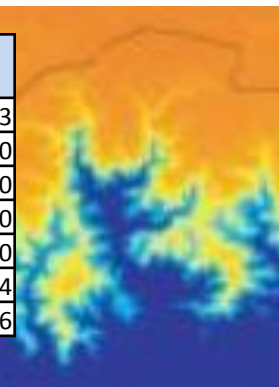
# Data Fusion – Small-Area Output

Aggregate data for administrative districts with tabulated population data and environmental characteristics

- Min/ max
- Mean
- Mode



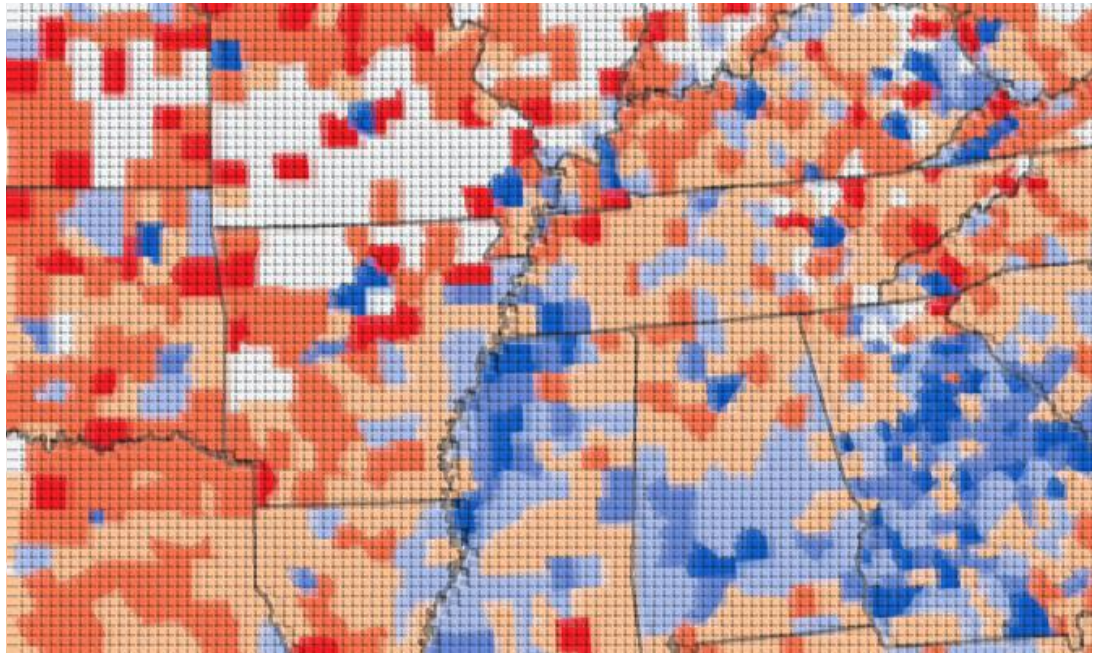
County ID	Mean Ann. Temp.	Max. Ann. Precip.	Rent, Rural	Rent, Urban	Own, Rural	Own, Urban	Vacant, Rural	Vacant, Urban
G17003100001	21.2	768	3129	1063	637	365	34	33
G17003100002	23.4	589	2949	1075	1469	717	0	0
G17003100003	24.3	867	3418	1589	1108	617	0	0
G17003100004	21.5	943	1882	425	202	142	123	0
G17003100005	24.1	867	2416	572	426	197	189	0
G17003100006	24.4	697	2560	934	950	563	220	14
G17003100007	25.6	701	2126	653	321	215	209	46



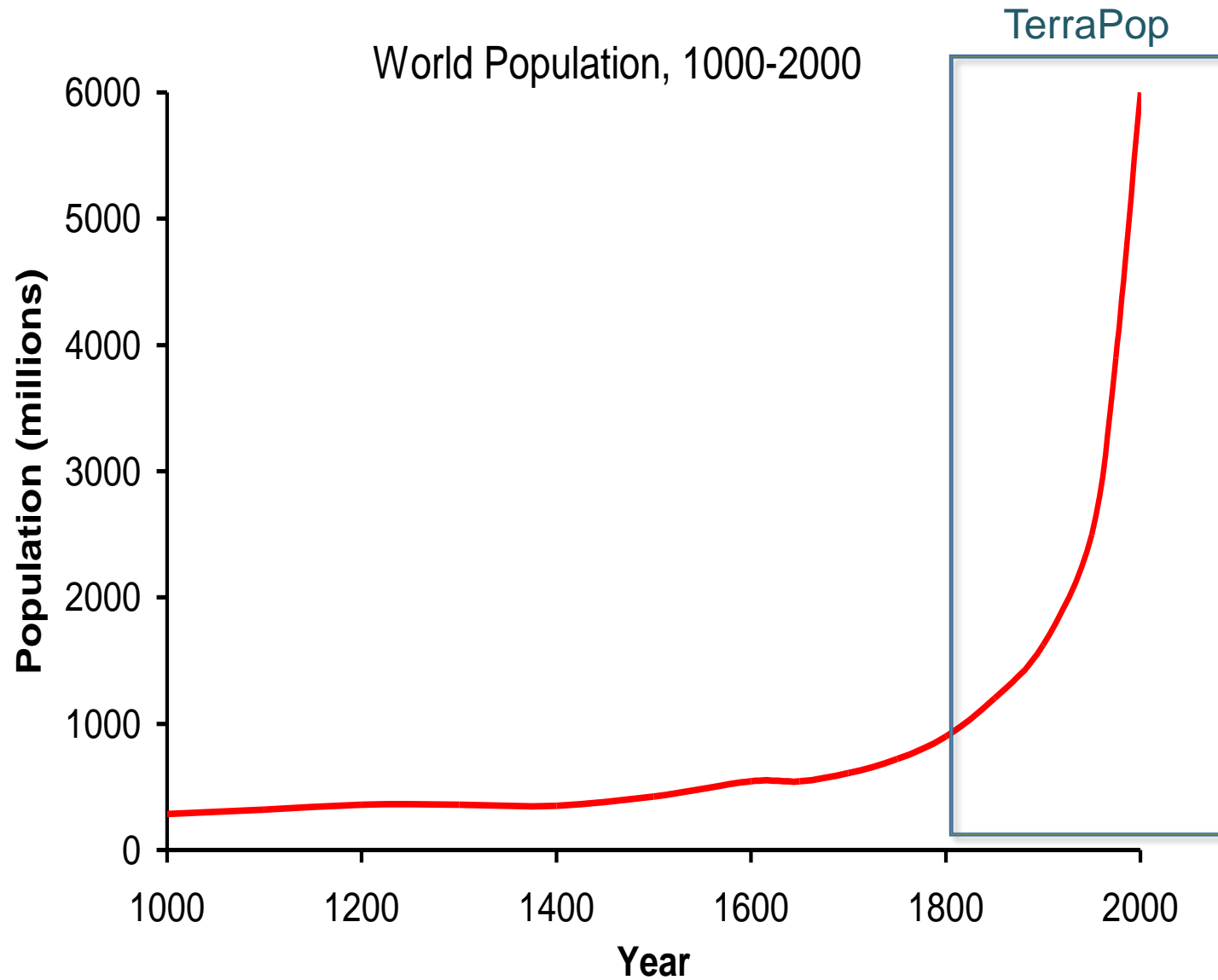
# Data Fusion – Gridded Output

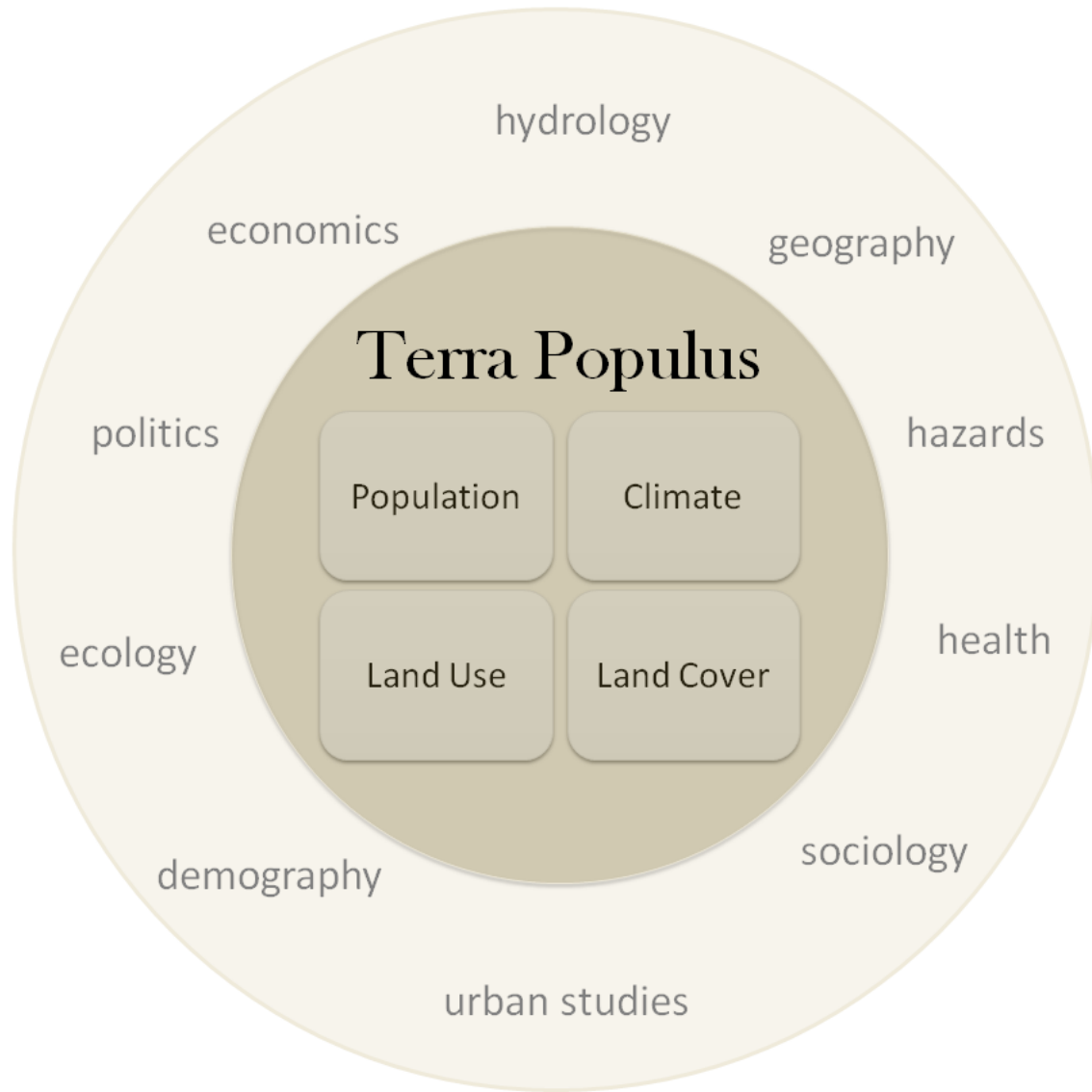
Gridded data with characteristics of population and environment

Generate  
population data in  
format compatible  
with environmental  
models



# The Temporal Dimension





# Sustainability

Create a sustainable organization that can guarantee preservation and access over multiple decades

- Organizational sustainability
- Financial sustainability
- Technological sustainability



